

Erdős-Szekeres without induction

Sergey Norin*

Yelena Yuditsky†

September 14, 2015

Abstract

Let $ES(n)$ be the minimal integer such that any set of $ES(n)$ points in the plane in general position contains n points in convex position. The problem of estimating $ES(n)$ was first formulated by Erdős and Szekeres [7], who proved that $ES(n) \leq \binom{2n-4}{n-2} + 1$. The current best upper bound, $\limsup_{n \rightarrow \infty} \frac{ES(n)}{\binom{2n-5}{n-2}} \leq \frac{29}{32}$, is due to Vlachos [15].

We improve this to

$$\limsup_{n \rightarrow \infty} \frac{ES(n)}{\binom{2n-5}{n-2}} \leq \frac{7}{8}.$$

1 Introduction

The following problem of Erdős and Szekeres has attracted considerable attention over the years.

The Erdős-Szekeres problem. For a positive integer $n \geq 3$, determine the smallest integer $ES(n)$ such that any set of at least $ES(n)$ points in general position in the plane contains n points that are the vertices of a convex n -gon.

The first bounds on $ES(n)$ were given by Erdős and Szekeres [7, 8], who have shown that

$$2^{n-2} + 1 \leq ES(n) \leq \binom{2n-4}{n-2} + 1.$$

*Department of Mathematics and Statistics, McGill University. Email: snorin@math.mcgill.ca. Supported by an NSERC grant 418520.

†School of Computer Science, McGill University. Email: yuditskyl@gmail.com.

They conjectured that the lower bound is tight. Their conjecture is verified for $n \leq 6$. It is trivial for $n = 3$; the proof for $n = 4$ was given by Esther Klein and can be found in [7]; the proof for $n = 5$, as stated in [7], was first given by Makai and can be found in [6]; the proof for $n = 6$ is due to Szekeres and Peters [12].

For the upper bound, although it is very far from the lower bound, little improvement was made over the last eighty years. The first such improvement was given by Chung and Graham [3], who improved the upper bound by one to $\binom{2n-4}{n-2}$. Shortly after, Kleitman and Pachter [4] showed that $ES(n) \leq \binom{2n-4}{n-2} + 7 - 2n$. Tóth and Valtr [13, 14] have shown that $ES(n) \leq \binom{2n-5}{n-2} + 1$. Finally, very recently Vlachos [15] have shown that $\limsup_{n \rightarrow \infty} \frac{ES(n)}{\binom{2n-5}{n-2}} \leq \frac{29}{32}$.

We further improve the upper bound as follows.

Theorem 1.1.

$$\limsup_{n \rightarrow \infty} \frac{ES(n)}{\binom{2n-5}{n-2}} \leq \frac{7}{8}.$$

Many variants of the Erdős-Szekeres problem were studied, and we refer the reader to surveys [1, 2, 9, 14].

Our proof of Theorem 1.1, unlike the original proofs of all the bounds listed above, doesn't use induction, but is based on assigning a word in two letter alphabet to each point and analyzing the possible sets of resulting words. We describe the technique in Section 2. We prove Theorem 1.1 in Section 3.

2 Encoding points

In this section we introduce the concepts and prove the basic results underlying the proof of Theorem 1.1.

For a positive integer m , let $[m] := \{1, 2, \dots, m\}$, and let $[m]^{(2)}$ denote the set of two element subsets of $[m]$.

Fix $f : [m]^{(2)} \rightarrow \mathbb{R}$. We say that a sequence (q_1, q_2, \dots, q_t) of elements of $[m]$, satisfying $q_1 < q_2 < \dots < q_t$ is a *t-cup with respect to f* if

$$f(q_1, q_2) \leq f(q_2, q_3) \leq \dots \leq f(q_{t-1}, q_t),$$

and this sequence is a *t-cap with respect to f* if

$$f(q_1, q_2) \geq f(q_2, q_3) \geq \dots \geq f(q_{t-1}, q_t).$$

(We write $f(x, y)$ instead of $f(\{x, y\})$ for brevity.) We say that a t -cup (or a t -cap) (q_1, q_2, \dots, q_t) *starts* at q_1 and *ends* in q_t .

We say that the function f is (k, l) -free if there exists no k -cup and no l -cap with respect to f . Let $ES'(k, l)$ denote the maximum m such that there exist a (k, l) -free $f : [m]^{(2)} \rightarrow \mathbb{R}$. One of the main byproducts of the construction introduced in this section is a new proof of the following result of Erdős and Szekeres.

Theorem 2.1 ([7]).

$$ES'(k+2, l+2) \leq \binom{k+l}{k}.$$

Our goal is to assign words to every element of $[m]$ which partially encode the structure of cups and caps ending and starting at this element. For every $i \in [m]$ we define four functions $\alpha^i, \gamma^i : \mathbb{N} \rightarrow \mathbb{R} \cup \{\pm\infty\}$ and $\beta^i, \delta^i : \mathbb{N} \rightarrow \mathbb{R} \cup \{\pm\infty\}$. The values of these functions at $t \in \mathbb{N}$ are defined as follows:

- Let $\alpha^i(t)$ be the minimum $f(j, i)$ such that there exists a $(t+1)$ -cup which ends with the pair (j, i) , i.e. a sequence $q_1 < q_2 < \dots < q_{t-1} < j < i$ such that

$$f(q_1, q_2) \leq \dots \leq f(q_{t-1}, j) \leq f(j, i).$$

If there is no $(t+1)$ -cup ending in i , let $\alpha^i(t) = +\infty$.

- Let $\beta^i(t)$ be the maximum $f(j, i)$ such that there exists a $(t+1)$ -cap which ends with the pair (j, i) . If there is no such cap then $\beta^i(t) = -\infty$.
- Let $\gamma^i(t)$ be the maximum $f(i, j)$ such that there exists a $(t+1)$ -cup which starts with the pair (i, j) , and if there is no such cup then $\gamma^i(t) = -\infty$.
- Finally, let $\delta^i(t)$ be the minimum $f(i, j)$ such that there exists a $(t+1)$ -cap which starts with the pair (i, j) , and if there is no such cap then $\delta^i(t) = +\infty$.

(The above functions encode the optimal values for further extending the cups and caps of given length which start or end at i .)

Assume now that the function f is $(k+2, l+2)$ -free. Let $\mathcal{L}^{k,l}$ denote the set of all words of length $k+l$ in the alphabet $\{\alpha, \beta\}$, which use exactly k

$i \backslash j$	1	2	3	4	5
1		2	4	-2	1
2	2		6	-6	-1
3	4	6		-9	-5
4	-2	-6	-9		7
5	1	-1	-5	7	

i	$\alpha^i(1)$	$\alpha^i(2)$	$\beta^i(1)$	$\beta^i(2)$	L_i
1	∞	∞	$-\infty$	$-\infty$	$\beta\beta\alpha\alpha$
2	2	∞	2	$-\infty$	$\beta\alpha\beta\alpha$
3	4	6	6	$-\infty$	$\beta\alpha\alpha\beta$
4	-9	∞	-2	-6	$\alpha\beta\beta\alpha$
5	-5	7	7	-1	$\alpha\alpha\beta\beta$

Figure 1: An example of a $(4, 4)$ -free function and corresponding L_i 's.

symbols “ α ” and l symbols “ β ”. We define a word $L_i \in \mathcal{L}^{k,l}$ for each $i \in [m]$ as follows. Sort the multiset $(\alpha^i(1), \alpha^i(2), \dots, \alpha^i(k), \beta^i(1), \beta^i(2), \dots, \beta^i(l))$ in increasing order, using a convention that when the elements repeat, then the values of the function α^i precede the values of the function β^i . The word L_i is obtained from the resulting sequence by replacing the values of α^i and β^i by the symbols α and β respectively. Figure 1 provides an example. (The first table lists the values of the function f and the second one the values of α^i , β^i and L_i .)

Let $\mathcal{R}^{k,l}$ denote the set of all words of length $k+l$ in the alphabet $\{\gamma, \delta\}$, which use exactly k symbols “ γ ” and l symbols “ δ ”. We define a word $R_i \in \mathcal{R}^{k,l}$, symmetrically to the word L_i , by sorting the values of γ_i and δ_i , using the convention that the δ symbols precede the γ symbols if the corresponding function values are equal.

We are now ready for the first result on $(k+2, l+2)$ -free functions.

Lemma 2.2. *Let $f : [m]^{(2)} \rightarrow \mathbb{R}$ be $(k+2, l+2)$ -free, and let $1 \leq i < j \leq m$. Then there exist $x \in [k], y \in [l]$, such that*

$$\alpha^i(x) > f(i, j) > \beta^i(y) \text{ and } \alpha^j(x) \leq f(i, j) \leq \beta^j(y). \quad (1)$$

Symmetrically, there exist $x' \in [k], y' \in [l]$, s.t.

$$\gamma^j(x') < f(i, j) < \delta^j(y') \text{ and } \gamma^i(x') \geq f(i, j) \geq \delta^i(y'). \quad (2)$$

Proof. By symmetry it suffices to prove (1). Let x be chosen minimum $\alpha^i(x) > f(i, j)$. Note that $x \leq k$, as $\alpha^i(k+1) = +\infty$. We have $\alpha^i(x-1) \leq f(i, j)$, or $x = 1$. Therefore, if $x \geq 2$ then there exists an x -cup (q_1, \dots, q_{x-1}, i) such that $f(q_{x-1}, i) \leq f(i, j)$. It follows that $(q_1, \dots, q_{x-1}, i, j)$ is an $(x+1)$ -cup. Thus $\alpha^i(x) \geq f(i, j)$.

Similarly, if y is chosen maximum such that $\beta^i(y) < f(i, j)$ then $f(i, j) \leq \beta^j(y)$. \square

The next corollary follows immediately from Lemma 2.2.

Corollary 2.3. *Let $f : [m]^{(2)} \rightarrow \mathbb{R}$ be $(k+2, l+2)$ -free, and let $1 \leq i < j \leq m$. Then $L_i \neq L_j$ and $R_i \neq R_j$.*

Note that Theorem 2.1 immediately follows from Corollary 2.3, as $|\mathcal{R}^{k,l}| = \binom{k+l}{k}$. The bijective proof of Theorem 2.1 presented in this section is due to the first author [10]. It is inspired by the beautiful argument used by Seidenberg [11] to give a bijective proof of the Erdős-Szekeres theorem on monotone subsequences of a sequence [7]. A different bijective proof was recently discovered by Moshkovitz and Shapira [5].

3 Proof of Theorem 1.1

We apply the framework introduced in the previous section to the Erdős-Szekeres problem as follows.

Let $P = \{p_1, p_2, \dots, p_m\}$ be a set of points in the plane in general position with Cartesian coordinates $(x_1, y_1), (x_2, y_2), \dots, (x_m, y_m)$, respectively. Assume without loss of generality that $x_1 < x_2 < \dots < x_m$. We define a function $f_P : [m]^{(2)} \rightarrow \mathbb{R}$ as follows. For $1 \leq i < j \leq m$, let $f_P(i, j)$ be the slope of the line that passes through the points i, j , i.e.

$$f_P(i, j) = \frac{y_j - y_i}{x_j - x_i}.$$

Assume that P does not contain n points that are the vertices of a convex n -gon. Thus f_P is (n, n) -free, as every n -cup (or n -cap) with respect to f_P

indexes the vertices of a convex n -gon. In the following, for simplicity, we will refer to the sets of points indexed by t -cups (respectively, t -caps) with respect to f_P as t -cups (respectively, t -caps) in P . We will refer to 3-cups and 3-caps as, simply, *cups* and *caps*.

We start by introducing two classical tools used to upper bound $ES(n)$ by Tóth and Valtr [14], and Erdős and Szekeres [7], respectively. The next theorem from [14], allows us to further restrict potential functions f_P .

Theorem 3.1 ([14]). *Let N be a positive integer such that every point set P with $|P| \geq N$ contains either a set of n points in convex position or a $(n-1)$ -cap. Then $ES(n) \leq N + 1$.*

Theorem 3.1 allows us to consider only point sets P such that f_P is $(n, n-1)$ -free.

The next lemma essentially appears in [7], but we include its proof for completeness.

Lemma 3.2 ([7]). *If f_P is $(n, n-1)$ -free then no point in P is simultaneously the end of an $(n-1)$ -cup and the beginning of an $(n-2)$ -cap. Symmetrically, no point p is the end of an $(n-2)$ -cap and the beginning of an $(n-1)$ -cup.*

Proof. Suppose for a contradiction that (q_1, \dots, q_{n-2}, q) is an $(n-1)$ -cup with respect to f_P and that $(q, q'_2, \dots, q'_{n-2})$ is an $(n-2)$ -cap. If $f(q_{n-2}, q) \leq f(q, q'_2)$ then $(q_1, \dots, q_{n-2}, q, q'_2)$ is an n -cup, and, otherwise, $(q_{n-2}, q, \dots, q'_{n-2})$ is an $(n-1)$ -cup, contradicting the fact that f_P is $(n, n-1)$ -free. \square

We'd like to extend the above lemma by showing that no point of P can be the end of an $(n-2)$ -cup and the beginning of both an $(n-2)$ -cap and $(n-1)$ -cup. Unfortunately, this statement is not necessarily true. Fortunately, it can be made true by removing a negligible proportion of points in P .

Let $Q \subseteq P$ be the set of points which serve as an end of an $(n-2)$ -cup and the beginning of both an $(n-2)$ -cap and $(n-1)$ -cup. For every point $q \in Q$, fix $V^q = (v_1^q, v_2^q, \dots, v_{n-3}^q, q)$ an $(n-2)$ -cup ending in q , $U^q = (q, u_2^q, \dots, u_{n-2}^q)$ an $(n-2)$ -cap starting in q , and $W^q := (q, w_2^q, w_3^q, \dots, w_{n-1}^q)$ an $(n-1)$ -cup starting in q . (See Figure 2.) We say that (V^q, U^q, W^q) is a Q -signature of q .

The following Claims 3.3 and 3.4 were inspired by [15], yet they appear not to follow directly from any of the results in [15].

Claim 3.3. *If (V, U, W) is a Q -signature of a point $q \in Q$, such that $U^q = (q, u_2, \dots, u_{n-2})$, $W^q = (q, w_2, \dots, w_{n-1})$, then $u_2 = w_{n-1}$.*

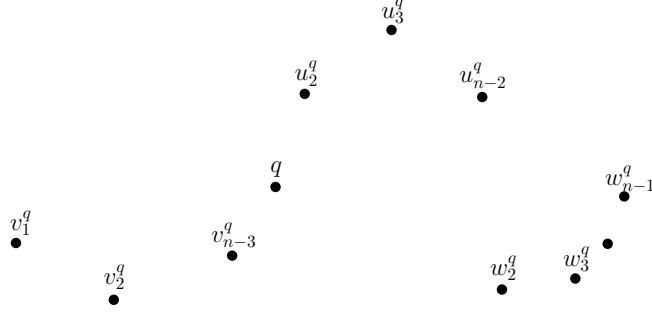


Figure 2: A point $q \in Q$.

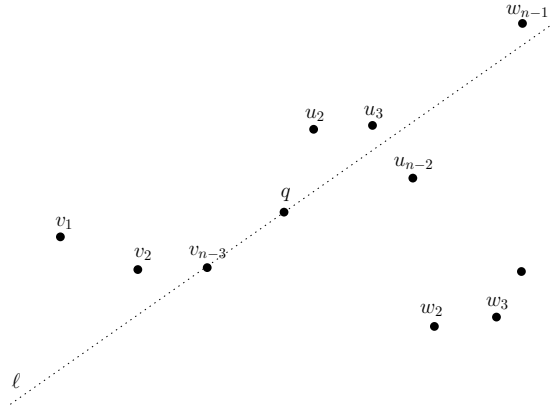


Figure 3: The point u_2 is to the left of w_{n-1} .

Proof. Suppose for a contradiction that $u_2 \neq w_{n-1}$. Let $V = (v_1, v_2, \dots, v_{n-3}, q)$ and let ℓ be the line passing through v_{n-3} and q .

If u_2 is below the line ℓ , then v_{n-3}, U is an $(n-1)$ -cap. Therefore u_2 is above ℓ . Similarly, if w_2 is above ℓ , then v_{n-3}, W is an n -cup, therefore w_2 is below ℓ .

Suppose that u_2 is to the left of w_{n-1} . If q, u_2, w_{n-1} is a cup then V, u_2, w_{n-1} is an n -cup. Otherwise, $W \cup \{u_2\}$ is a vertex set of a convex n -gon. (See Figure 3.)

Therefore u_2 is to the right to w_{n-1} . Suppose that w_{n-2}, u_2, u_3 is a cap. Then w_{n-3}, w_{n-2}, u_2 is a cup, as otherwise $w_{n-3}, w_{n-2}, u_2, \dots, u_{n-2}$ is an $(n-1)$ -cap. If q, w_{n-1}, u_2 is a cap, then $q, w_{n-1}, u_2, w_{n-2}, w_{n-3}, \dots, w_2$ is a vertex set convex n -gon. Therefore q, w_{n-1}, u_2 is a cup, implying that w_{n-1}, u_2, u_3 is

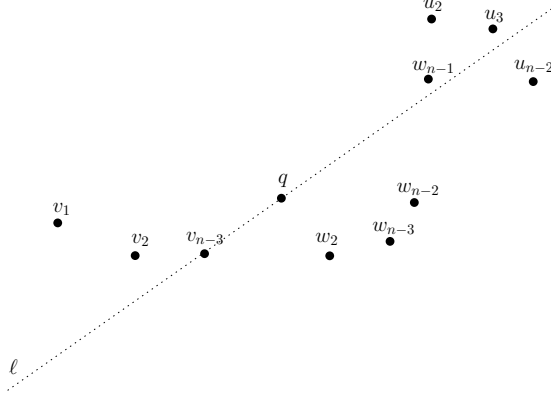


Figure 4: The point u_2 is to the right of w_{n-1} .

a cap. If w_{n-2}, w_{n-1}, u_2 is a cap, then $w_{n-2}, w_{n-1}, u_2, u_3, \dots, u_{n-2}$ is an $(n-1)$ -cap. Thus w_{n-2}, w_{n-1}, u_2 is a cup, therefore W, u_2 is an n -cup. (See Figure 4.)

It remains to consider the subcase when w_{n-2}, u_2, u_3 is a cup. In this case v_{n-3}, q, w_{n-2} is a cup, and thus V, w_{n-2}, w_{n-1} is an n -cup, yielding the final contradiction. \square

We say that $q' \in P$ is a *mate* of a point $q \in Q$ if there exist a Q -signature (V, U, W) of q with points of U and W labelled as in Claim 3.3, such that $q' = u_2 = w_{n-1}$. Let Q' be the set of all mates of points in Q . By Claim 3.3 every point in Q has at least one mate.

Claim 3.4. $|Q'| \leq |P|/(n-2)$.

Proof. Consider $q'_1, q'_2 \in Q'$. Let q'_1 and q'_2 be the mates of q_1 and q_2 , respectively. Let $(V^{q_1}, U^{q_1}, W^{q_1}), (V^{q_2}, U^{q_2}, W^{q_2})$ be Q -signatures of q_1 and q_2 with points labelled as in the definition such that $q'_1 = u_2^{q_1} = w_{n-1}^{q_1}$ and $q'_2 = u_2^{q_2} = w_{n-1}^{q_2}$. We assume without loss of generality that q'_1 is to the left of q'_2 . We will show that q_2 is to the right of q'_1 . This would imply the claim, as it will follow that the points of W^{q_2} lie to the right of q'_1 and to the left of q_2 , implying that the indices of any two points in Q' differ by at least $n-2$.

Assume to the contrary that q_2 is to the left of q'_1 . We claim that all the points of P to the right of q_2 lie below the line that passes through q_2 and q'_2 . Indeed, suppose that r is a point above this line. If r is to the left of q'_2 then we find a convex n -gon in P , using W^{q_2} and r . If r is to the

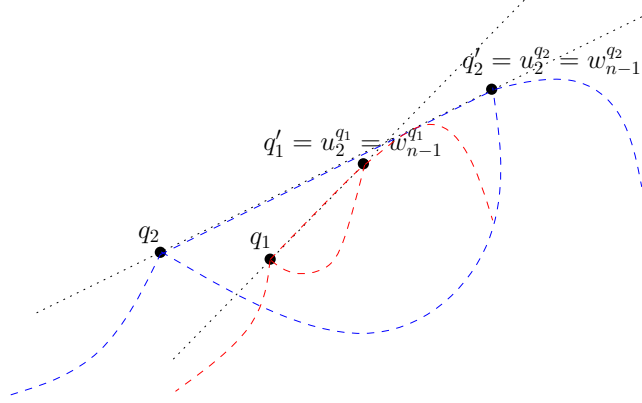


Figure 5: A possible arrangement of points in Claim 3.4.

right of q'_2 then we find a n -cup using V^{q_2}, q'_2 and r . (Note that V^{q_2}, q'_2 is an $(n-1)$ -cup, because otherwise $v_{n-3}^{q_2}, U^{q_2}$ is an $(n-1)$ -cap.). Analogously, all the points of P to the right of q_1 must be below the line that goes through q_1, q'_1 . (See Figure 5 which illustrates a possible arrangement of points in this case.) Thus $q_1, q'_1, q'_2, u_3^{q_2}, u_4^{q_2}, \dots, u_{n-2}^{q_2}$ is an $(n-1)$ -cap, a contradiction. \square

As our goal is to prove $m \leq (\frac{7}{8} + o(1)) \binom{2n-5}{n-2}$, by Claim 3.4 we may consider $P - Q'$ instead of P , and thus assume that no point of P is simultaneously an end of an $(n-2)$ -cup and the beginning of both an $(n-2)$ -cap and $(n-1)$ -cup.

It is time to utilize the encoding of points introduced in Section 2. Let $L_1, \dots, L_m \in \mathcal{L}^{n-2, n-3}$ and $R_1, \dots, R_m \in \mathcal{R}^{n-2, n-3}$ be the words associated with points of P via the function f_P defined in Section 2. We will write \mathcal{L} and \mathcal{R} instead of $\mathcal{L}^{n-2, n-3}$ and $\mathcal{R}^{n-2, n-3}$, for brevity. For a pair of words w_1 and w_2 , let $\mathcal{L}(w_1 * w_2)$ denote the set of words in \mathcal{L} which start with w_1 and end with w_2 . We write $\mathcal{L}(*w_2)$ and $\mathcal{L}(w_1*)$, instead of $\mathcal{L}(w_1 * w_2)$, if w_1 (respectively, w_2) is the null word. The sets $\mathcal{R}(w_1 * w_2)$ are defined analogously. The next claim follows immediately from the definitions of L_i and R_i .

Claim 3.5. *For $i \in [m]$ the following statements hold*

- p_i is the end of an $(n-1)$ -cup if and only if $L_i \in \mathcal{L}(*\beta)$,
- p_i is the end of an $(n-2)$ -cap if and only if $L_i \in \mathcal{L}(\alpha*)$,

- p_i is the beginning of an $(n - 1)$ -cup if and only if $R_i \in \mathcal{R}(*\gamma)$,
- p_i is the beginning of an $(n - 2)$ -cap if and only if $R_i \in \mathcal{R}(\delta*)$, and
- p_i is not an end of any $(n - 2)$ -cup if and only if $L_i \in \mathcal{L}(*\alpha\alpha)$.

Claim 3.5, Lemma 3.2 and our assumptions on P , based on Claims 3.3 and 3.4, imply the following.

Claim 3.6. *If $R_i \in \mathcal{R}(\delta * \gamma)$ for some $i \in [m]$, then $L_i \in \mathcal{L}(\beta * \alpha\alpha)$.*

We are now ready to finish the proof of Theorem 1.1. Let $\mathcal{S} = \{R_1, \dots, R_m\}$. By Claim 3.6 and Corollary 2.3, we have

$$|\mathcal{S} \cap \mathcal{R}(\delta * \gamma)| \leq |\mathcal{L}(\beta * \alpha\alpha)| = \left(\frac{1}{8} + o(1)\right) \binom{2n-5}{n-2}.$$

Thus, by Corollary 2.3,

$$m = |\mathcal{S}| \leq |\mathcal{S} \cap \mathcal{R}(\delta * \gamma)| + |\mathcal{R} - \mathcal{R}(\delta * \gamma)| \leq \left(\frac{7}{8} + o(1)\right) \binom{2n-5}{n-2},$$

as desired.

References

- [1] I. Bárány and Gy. Károlyi, Problems and results around the Erdős–Szekeres theorem, *Japanese Conference on Discrete and Computational Geometry* (2001), 91-105.
- [2] P. Brass, W. Moser, and J. Pach, Convex polygons and the Erdős–Szekeres problem, Chapter 8.2 in *Research problems in discrete geometry*, Springer, (2005).
- [3] F. R. K. Chung and R. L. Graham, Forced convex n -gons in the plane, *Discrete and Computational Geometry* **19** (1998), 367-371.
- [4] D. J. Kleitman and L. Pachter, Finding convex sets among points in the plane, *Discrete and Computational Geometry* **19** (1998), 405-410.
- [5] G. Moshkovitz and A. Shapira, Ramsey Theory, Integer Partitions and a New Proof of the Erdős-Szekeres Theorem, *Advances in Mathematics* **262** (2014), 1107-1129.

- [6] J.D. Kalbfleisch, J.G. Kalbfleisch and R.G. Stanton, A combinatorial problem on convex regions, *Proc. Louisiana Conf. Combinatorics, Graph Theory and Computing* (1970), 180-188.
- [7] P. Erdős and G. Szekeres, A combinatorial problem in geometry, *Compositio Mathematica* **2** (1935), 463-470.
- [8] P. Erdős and G. Szekeres, On some extremum problems in elementary geometry, *Ann. Univ. Sci. Budapest. Etsv Sect. Math.* **34** (1960-1961), 53-62.
- [9] W. Morris and V. Soltan, The Erdős-Szekeres problem on points in convex position a survey. *Bull. Amer. Math. Soc.* **37** (2000), 437-458.
- [10] S. Norin, Erdős-Szekeres for first differences, <http://mathoverflow.net/q/91128>, (2012).
- [11] A. Seidenberg, A simple proof of a theorem of Erdős and Szekeres., *J. London Math. Soc.* **34** (1959), 352.
- [12] G. Szekeres and L. Peters, Computer solution to the 17-point Erdős-Szekeres problem, *ANZIAM Journal* **48** (2006), 151-164.
- [13] G. Tóth and P. Valtr, Note on the Erdős-Szekeres theorem, *Discrete and Computational Geometry* **19** (1998), 457-459.
- [14] G. Tóth and P. Valtr, The Erdős-Szekeres theorem: Upper bounds and related results, *Combinatorial and Computational Geometry* (J.E. Goodman et al., eds.), Publ. M.S.R.I. **52** (2006) 557-568.
- [15] G. Vlachos, On a conjecture of Erdős and Szekeres, <http://arxiv.org/abs/1505.07549>, (2015).